HUMAN MUTATION 26(3), 262–270, 2005

METHODS

Comparison of Yield and Genotyping Performance of Multiple Displacement Amplification and OmniPlex™ Whole Genome Amplified DNA Generated From Multiple DNA Sources

Andrew W. Bergen,[1]* Kashif A. Haque,[2,3] Ying Qi,[2,3] Michael B. Beerman,[2,3] Montserrat Garcia-Closas,[1] Nathaniel Rothman,[1] and Stephen J. Chanock[1,2,4]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland; [2]Core Genotyping Facility, National Cancer Institute, National Institutes of Health, Gaithersburg, Maryland; [3]Intramural Research Support Program, SAIC-Frederick, NCI-FCRDC, Frederick, Maryland; [4]Section on Genomic Variation, Pediatric Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland

Communicated by Ann-Christine Syvänen

The promise of whole genome amplification (WGA) is that genomic DNA (gDNA) quantity will not limit molecular genetic analyses. Multiple displacement amplification (MDA) and the OmniPlex™ PCR-based WGA protocols were evaluated using 4 and 5 ng of input gDNA from 60 gDNA samples from three tissue sources (mouthwash, buffy coat, and lymphoblast). WGA DNA (wgaDNA) yield and genotyping performance were evaluated using genotypes determined from gDNA and wgaDNA using the AmpF*l*STR® Identifiler® assay and N = 49 TaqMan® SNP assays. Short tandem repeat (STR) and SNP genotyping completion and concordance rates were significantly reduced with wgaDNA from all WGA methods compared with gDNA. OmniPlex wgaDNA exhibited a greater reduction in genotyping performance than MDA wgaDNA. Reduced wgaDNA genotyping performance was due to allelic (all protocols) and locus (OmniPlex) amplification bias leading to heterozygote and locus dropout, respectively, and %GC sequence content (%GC) was significantly correlated with TaqMan assay performance. Lymphoblast wgaDNA exhibited higher yield (OmniPlex), buffy coat wgaDNA exhibited higher STR genotyping completion (MDA), whereas mouthwash wgaDNA exhibited higher SNP genotyping discordance (MDA). Genotyping of wgaDNA generated from ≤5 ng gDNA, e.g., from archaeological, forensic, prenatal diagnostic, or pathology samples, may require additional genotyping validation with gDNA and/or more sophisticated analysis of genotypes incorporating observed reductions in genotyping performance. Hum Mutat 26(3), 262–270, 2005. © 2005 Wiley-Liss, Inc.

KEY WORDS: whole genome amplification; multiple displacement amplification; GenomiPhi™; REPLI-g™ 2500S; OmniPlex™; AmpF*l*STR® Identifiler® Assay; TaqMan® Assay

INTRODUCTION

Genomic DNA (gDNA) sample quantity is a limiting resource in some genetic studies. Although high-throughput genotyping technologies now use less DNA per assay than in the past, the need to conserve valuable DNA samples is still a concern for investigators, biorepositories, and genotyping laboratories as the number of potential genotyping assays grows. Methods to faithfully amplify genomic DNA from minute quantities of starting material represent a solution to the challenge of ensuring adequate amounts of DNA of acceptable quality for high-throughput genetic analyses. PCR has enabled the amplification and analysis of specific loci within the genome [Mullis et al., 1986], but single-locus PCR is not efficient enough for multilocus studies when gDNA is limited. Whole genome amplification (WGA) can provide the assurance that gDNA quantity will not limit design and execution of molecular epidemiologic, clinical diagnostic, and forensic studies. Specifically, precious gDNA samples can be amplified using minute amounts of template, the whole genome

The Supplementary Material referred to in this article can be accessed at http://www.interscience.wiley.com/jpages/1059-7794/suppmat.

Received 6 December 2004; accepted revised manuscript 15 March 2005.

*Correspondence to: A.W. Bergen, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, 6120 Executive Boulevard, Bethesda, MD. 20892-7236. E-mail: bergena@mail.nih.gov

Grant sponsor: National Cancer Institute; Grant sponsor: National Institutes of Health; Grant number: NO1-CO-12400.

DOI 10.1002/humu.20213
Published online in Wiley InterScience (www.interscience.wiley.com).

amplified DNA (wgaDNA) used for high-throughput genotyping and sequencing studies, and the remaining gDNA stored for validation or future applications.

During the past dozen years, PCR-based methods of WGA have been used to amplify whole genomes: degenerate oligonucleotide primed amplification (DOP) [Telenius et al., 1992] and primer-extension preamplification (PEP) [Zhang et al., 1992]. However, these two methods (and modifications thereof) can result in uneven amplification of both loci and alleles [Cheung and Nelson, 1996; Dean et al., 2002; Dietmaier et al., 1999; Paunio et al., 1996; Wells et al., 1999]. Recently, novel methods for the amplification of whole genomes have been presented. Multiple displacement amplification (MDA) uses the highly processive bacteriophage ϕ29 DNA polymerase and degenerate hexamers to amplify DNA [Dean et al., 2001, 2002]. This method has been successful in amplifying minute amounts of gDNA from whole blood, dried blood, buccal cells, cultured cells, and buffy coats, resulting in a ~10,000-fold increase of gDNA [Dean et al., 2002; Hosono et al., 2003]. Another WGA method, introduced by Rubicon Genomics (Ann Arbor, MI; www.rubicongenomics.com), employs OmniPlex™ libraries of 200–2,000-bp fragments created by random chemical cleavage of gDNA, followed by ligation of adaptor sequences to both ends and PCR amplification [Langmore, 2002]. This fragmentation/ligation/PCR-based method amplifies the entire genome with less than 0.043% locus dropout, based on analysis of a whole genome SNP linkage panel [Barker et al., 2004], thus overcoming a serious limitation of previous PCR-based WGA methods similar to OmniPlex, which have exhibited a mean short tandem repeat (STR) and STS locus amplification failure rate of 21% [Klein et al., 1999; Lucito et al., 1998; Tanabe et al., 2003].

This study is a comparative analysis of three commercially available WGA protocols (two MDA-based and one PCR-based) using N = 60 DNA from three tissue sources commonly collected in molecular epidemiological studies: lymphoblasts, mouthwash cell pellets, and buffy coat samples. wgaDNA was quantified to determine the yield and composition of the wgaDNA produced by each protocol. Both gDNA and wgaDNA were genotyped in duplicate at a total of 65 STR and SNP loci using the AmpF*l*STR® Identifiler® (Applied Biosystems, Foster City, CA; www.appliedbiosystems.com) and TaqMan® (Applied Biosystems, Foster City, CA; www.appliedbiosystems.com) assays. Genotype completion and discordance rates of gDNA and wgaDNA were calculated and compared to assess genotyping performance of wgaDNA to guide the use of wgaDNA in molecular epidemiology studies.

MATERIALS AND METHODS

Samples

gDNA obtained from lymphoblasts (N = 20; Coriell Cell Repositories, Camden, NJ, http://locus.umdnj.edu/ccr), and matched gDNA samples extracted from mouthwash cell pellets (hereafter, "mouthwash") and buffy coat samples (hereafter, "buffy coat") of 20 individuals was used to evaluate the three WGA protocols (total N = 60 gDNA samples). We recruited 21 volunteers (15 female) aged 42–65 years from the Division of Cancer Epidemiology and Genetics of the National Cancer Institute (NCI), in order to obtain matched buffy coat and mouthwash gDNA samples. The NCI Institutional Review Board (IRB) approved the studies and all participants provided signed informed consent. Volunteers were asked not to brush, eat, or drink 1 hr prior to sample collection, and not to swallow saliva during the collection process. We collected a 10-mL blood sample in a lavender-top Vacutainer™ (Becton, Dickinson and Company, Franklin Lakes, NJ; www.bd.com) tube and mouthwash samples with 10 mL of mouthwash solution (Scope® (Procter & Gamble Company, Cincinnati, OH; www.pg.com)). Volunteers vigorously swished the mouthwash around in his/her mouth for 45 sec, and deposited the mouthwash back into the collection container. Blood samples were transported to the laboratory at room temperature (RT) for processing, and upon arrival to the laboratory, refrigerated at 4°C until processing the next day using a standard processing protocol to obtain plasma, red blood cells, and buffy coats, which were frozen immediately at –80°C. Mouthwash samples were transported to the laboratory at RT and processed after 3 days at RT to mimic the typical delay in processing when samples are collected by mail. Mouthwash samples were transferred to a 50-mL conical tube, centrifuged at 1500 *g* for 15 min, the supernatant was discarded, and the cell pellet was resuspended in 3 mL of TE buffer solution, mixed well, aliquotted equally into two cryovials of volume and stored until DNA extraction at –80°C. DNA was extracted from buffy coat and mouthwash cell pellets using a phenol-chloroform extraction protocol [Garcia-Closas et al., 2001].

Whole Genome Amplification Reactions

Each gDNA was quantified by ultraviolet (UV) spectroscopy, the PicoGreen® assay (Molecular Probes, Eugene, OR), and a real-time TaqMan® assay specific to human DNA [Haque et al., 2003]. Fourteen out of the 20 lymphoblast gDNAs were replicated for a total of N = 34 lymphoblast gDNA samples subjected to WGA to increase statistical power to detect genotyping discordance in lymphoblast gDNA analyses. A total of 4.0 ng of gDNA was used as template for the two MDA protocols, GenomiPhi™ DNA Amplification Kit (Amersham Biosciences, Piscataway, NJ), hereafter referred to as "MDA1," and REPLI-g™ 2500S Whole Genome Amplification Kit (Molecular Staging, New Haven, CT), hereafter referred to as "MDA2". Amplification reaction conditions for MDA1 follow those given by the manufacturer, with the exception that DNA was not denatured prior to reaction setup as it was observed that elimination of the denaturation step at 95°C did not negatively affect the yield of the MDA1 protocol (data not shown). MDA2 protocol conditions followed those given by the manufacturer, which included a chemical denaturation with KOH. Reaction volumes for the MDA2 protocol were proportionally reduced from the 1× reaction to a 0.4× reaction (20.0 μL) to match the volume of the MDA1 protocol, after we observed that reducing reaction volumes proportionally reduced wgaDNA yield (data not shown). A total of 30.0 ng of gDNA was sent to Rubicon Genomics, their proprietary OmniPlex WGA protocol (hereafter, "FLP" for fragmentation, adaptor-ligation PCR) was performed on 5.0 ng of gDNA by Rubicon Genomics and wgaDNA returned to NCI. Each gDNA sample was subjected to each WGA protocol once, along with no gDNA template controls (NTC) reactions in triplicate for the MDA protocols, and singly for FLP. These gDNA input masses were those recommended by WGA protocol inventors and vendors through 2003 (John P. Langmore, personal communication, September 4, 2003) [Dean et al., 2002; Lizardi, 2000].

wgaDNA obtained from all three protocols was precipitated using a standard ethanol precipitation protocol and resuspended in 100 μL of TE buffer. The mass of double-stranded wgaDNA was estimated using the PicoGreen assay. The mass of single-stranded wgaDNA was estimated using both OD_{260} and PicoGreen assay measurements according to the following transformation from the observed PicoGreen and OD_{260} measurements: the expected OD_{260} measurement for the double-stranded DNA (dsDNA) concentration

estimated by the observed PicoGreen measurement was calculated assuming 50-ng dsDNA/μL/A°_{260}, the calculated expected OD_{260} measurement for the dsDNA concentration was subtracted from the observed Optical Density at 260 nm (OD_{260}) measurement, and the difference was used to calculate the expected single-stranded DNA (ssDNA) concentration in the sample assuming 33 ng ssDNA/μl/A°_{260}. The total wgaDNA mass was obtained from the sum of the dsDNA and ssDNA masses as estimated using the OD_{260} and PicoGreen assay measurements. Wilcoxon's rank sum test was used to evaluate differences in wgaDNA yield distributions.

AmpF*l*STR Identifiler Assay

A total of 300 pg of dsDNA (both gDNA and wgaDNA, as determined by PicoGreen) was used as template DNA for the AmpF*l*STR Identifiler assay, which attempts to produce genotypes for N = 15 STR loci and for the Amelogenin (AMEL) locus for each DNA sample. The reaction conditions were: 2.10 μL of PCR Reaction Mix, 1.10 μL PCR Identifiler Primer Set, 1 U AmpliTaq Gold™, and add a sufficient quantity of (q.s.) water to 5.0 μL. Cycling conditions were: 95°C for 11 min, 28 cycles of 94°C for 1 min, 57°C for 1 min, and 72°C for 1 min, followed by 60°C for 60 min. The reaction was performed in duplicate and 1.0 μL of each reaction was run on an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). Scoring of STR alleles was automatically determined using ABI Prism® GeneMapper™ v3.0 software (Applied Biosystems, Foster City, CA), and each duplicate was independently genotyped.

The minimum GeneMapper v3.0 software thresholds (default parameter settings) are 50 relative fluorescence units (RFUs) for detection of STR allele peaks, 200 and 100 RFUs for "low signal strength" peaks for homozygous and heterozygous microsatellite genotypes, and a peak height ratio of allele 2 peak height/allele 1 peak height (PHR) of 50%. For the purpose of this analysis, a completed STR genotype and a no amplification genotype failure was considered to be an instance and no instance, respectively, of an allele peak above the minimum RFU threshold within the allele bin defined by the GeneMapper v3.0 software for the AmpF*l*STR Identifiler assay. The wgaDNA discordance rate was calculated to be the number of instances in which a wgaDNA STR genotype differed from the gDNA STR genotype. STR genotype completion, no amplification, and discordance rates were calculated with minimum signal strength thresholds of 50, 100, 150, 200, and 250 RFUs. Peak height ratio distributions at a signal strength threshold of ≥ 50 RFUs were evaluated for normality and differences between distributions evaluated using Wilcoxon's rank sum test. Genotype quality (GQ) scores (range 0.0–0.90) were assigned to each genotype by GeneMapper v 3.0 software. GQ scores <0.40 were flagged for manual editing and those that were edited successfully were assigned a GQ score of 1.0 by default by the software, while GQ scores of <0.25 were considered, in this analysis, as genotype failures. GQ scores of 1.0 were reassigned to a category of 0.25–0.39 for descriptive purposes in this analysis, along with the two other observed GQ score categories of scores of 0.40–0.45 and 0.80–0.90, and distributions were evaluated using contingency table analysis.

TaqMan SNP Assay Genotyping

A total of 4.0 ng of dsDNA (both gDNA and wgaDNA, as determined by PicoGreen) was used as template for SNP genotyping using the TaqMan assay at N = 49 SNPs (Supplementary Table S1, available online at http://www.interscience.wiley.com/jpages/1059-7794/suppmat), designed using the Assay-By-Design service (Applied Biosystems, Foster City, CA), optimized, and genotypes validated against resequencing data from 102 individuals in the SNP500Cancer program [Packer et al., 2004]. Assays were chosen from the SNP500Cancer database to be distributed on each chromosome arm excluding the acrocentric p-arms and the entire Y-chromosome, to exhibit a range of minor allele frequencies from 1 to 49.5%, and a range of %GC sequence content (hereafter "%GC") at the TaqMan assay PCR amplicon (mean amplicon size: 81 bp, median: 77 bp, range: 60–122 bp; mean %GC: 49.9%, median: 49.6%, range: 31.1–68.0%). Reaction and cycle conditions for all SNP assays were as follows: 1 × Universal Master Mix, 900 nM primers, 200 nM FAM-probe, 200 nM VIC-probe, and q.s. water for a final volume of 5.0 μL; and 50°C for 2 min, 92°C for 10 min, 50 cycles of 92°C for 30 sec, and 60°C for 1 min, performed in a 384-well plate. Details on all TaqMan SNP assay protocols used in this study are publicly available at the SNP500Cancer database website. gDNA and wgaDNA were amplified simultaneously in duplicate, and genotypes were determined. For quality control purposes, each assay replicate included three gDNA genotype controls (Allele 1, Allele 2, and a heterozygous sample, three replicates each) and three negative controls (water) to aid in the scoring of genotype, and NTC (no amplification) clusters. Fluorescence was detected using an ABI 7900HT Sequence Detection System and genotype clusters were manually scored using Sequence Detection Software 2.0 (Applied Biosystems, Foster City, CA). TaqMan SNP genotype attempts were classified as completed (fluorescent data point observed within 1 of the 2 homozygote clusters and in the heterozygote cluster), undetermined (fluorescent data point in 0 of the 2 homozygote clusters, the heterozygote cluster, or the NTC clusters) or no amplification (fluorescent data point in the NTC cluster).

SNP genotype completion, undetermined, and no amplification rates were calculated by dividing the number of completed, undetermined, and no amplification genotypes by the number of attempted genotypes; SNP genotype discordance rate was calculated by dividing the number of discordant genotypes by the number of completed genotypes. The wgaDNA discordance rate was defined as the number of instances in which a wgaDNA SNP genotype differed from the scored gDNA SNP genotype. The gDNA discordance rate was defined as the number of instances where the gDNA replicate genotypes were not concordant. For such discordances, the correct gDNA SNP genotype was assumed to be that given at the SNP500Cancer database for the particular SNP assay. Differences in completion, undetermined genotype, no amplification, and discordance rates were evaluated using contingency table significance testing. Spearman's rho sequential rank correlation analysis of the relationship between regional and amplicon %GC and gDNA and wgaDNA discordant and undetermined genotype counts was performed.

Analysis Software

Data was managed and/or analyzed using Microsoft Access and Excel (Microsoft, Redmond, WA), SISA (http://home.clara.net/sisa), SPSS (Chicago, IL), and SAS (SAS Institute, Cary, NC) software.

RESULTS

wgaDNA Yield and Composition

The median fold increase in wgaDNA mass over input DNA mass ranged from ~1,500-fold (MDA1) to ~5,100-fold (MDA2), with FLP exhibiting an intermediate increase of ~1,700-fold (Table 1). The methods differ in the mass of wgaDNA produced, due primarily to predetermined experimental WGA reaction volume, as the variance within each WGA protocol is modest

(e.g., mean coefficient of variation [CV] is 16.4%, 9.2%, and 12.9% for MDA1, MDA2, and FLP, respectively). The proportion of ssDNA produced by the three protocols is estimated to constitute ~8, 20, and 66% of the MDA1, MDA2, and FLP wgaDNA samples, respectively; the proportions significantly differ from each other (Table 1).

AmpF*l*STR Identifiler Profiling

To determine how accurately WGA reproduces polymorphic, repetitive sequences, each wgaDNA sample was genotyped by the AmpF*l*STR Identifiler assay and the resulting genotypes compared to genotypes obtained from gDNA (Table 2; Supplementary Tables S2 and S3). gDNA exhibited a genotype concordance rate of 100% at all RFU thresholds, with a slight drop in completion from 50 to 250 RFUs, reflecting a small number of low signal strength peaks. gDNA exhibited significantly greater STR genotype completion and concordance rates compared to all WGA protocols and wgaDNA tissue sources at all RFU thresholds (Table 2). There were no significant differences between the MDA protocols with respect to STR genotyping performance. The MDA protocols resulted in significantly higher STR genotyping completion and concordance rates compared to the FLP protocol at all RFU thresholds (Table 2). Six and two STR loci accounted for the majority of discordant and no amplification STR genotypes, respectively, following MDA, while seven and two STR loci accounted for the majority of discordant and no amplification STR genotypes for FLP, respectively (Supplementary Table S2).

gDNA STR genotype quality exceeded wgaDNA STR genotype quality (Supplementary Table S3). For example, gDNA was observed to produce only concordant STR genotypes, while wgaDNA produced both concordant and discordant STR genotypes, although FLP wgaDNA produced substantially more discordant STR genotypes than did MDA wgaDNA. PHRs and GQ scores differed in a complex way among gDNA, MDA, and FLP STR genotypes, though some differences are similar to the differences observed in completion and concordance rates for STR genotypes (Table 2). For example, PHRs observed with gDNA were similar to PHRs previously reported [Moretti et al., 2001], and PHRs did not differ within MDA protocols, or between gDNA and MDA concordant heterozygote genotypes. However, FLP concordant genotype PHRs were significantly different than gDNA and MDA concordant genotype PHRs, and FLP discordant genotype PHRs were significantly different from MDA discordant genotype PHRs.

For FLP wgaDNA, stutter peaks were observed with a fluorescence amplitude strong enough to be called as one of the two strongest alleles at several AmpF*l*STR Identifiler STR loci. Stutter peaks observed with MDA wgaDNA had low peak amplitudes similar to those of gDNA, consistent with the high processivity of ϕ29 DNA polymerase [Blanco et al., 1989]. While FLP wgaDNA stutter peaks contributed in part to the higher FLP discordance rate (1.9% of discordant FLP STR genotypes), the majority of both FLP and MDA discordant genotypes were due to wgaDNA samples exhibiting homozygote genotypes that were observed to be genotyped as heterozygotes, or as homozygotes of the other allele, when genotyping the corresponding gDNA sample. For most of those discordant genotypes in which the gDNA genotype was heterozygous, the wgaDNA was scored as homozygous for the shorter allele and a second allele peak was observed with the correct length, but with amplitude below the default PHR threshold. The mean PHRs for such discordant

TABLE 1. Summary of wgaDNA Yield

WGA protocol[a]	gDNA Tissue source[b,c]	Reaction concentration [X]	gDNA template [ng]	Median wgaDNA yield [μg][d]	Min. wgaDNA yield [μg]	Max. wgaDNA yield [μg]	~Fold increase	Median ssDNA [%][c,e,f]	Min. ssDNA [%]	Max. ssDNA [%]
MDA1	Mouthwash	1.0	4	5.97	4.63	8.80	1491	6.95	0.00	17.06
	Buffy coat	1.0	4	6.43	4.15	8.79	1608	7.62	2.44	19.72
	Lymphoblast	1.0	4	6.12	4.96	8.56	1530	9.37	4.91	17.13
MDA2	Mouthwash	0.4	4	20.19	14.93	21.38	5046	18.55	7.11	34.33
	Buffy coat	0.4	4	20.79	11.47	22.63	5198	23.30	6.79	40.30
	Lymphoblast	0.4	4	20.47	17.94	22.35	5116	16.58	6.44	32.23
FLP	Mouthwash	Not given	5	8.70	7.48	9.82	1740	65.79[g]	64.09	67.06
	Buffy coat	Not given	5	9.11[h]	1.18	10.29	1821	65.50[g]	63.82	67.70
	Lymphoblast	Not given	5	8.35	7.56	9.76	1670	66.46[g]	63.60	67.03

[a]Protocols: MDA1 = Amersham GenomiPhi™, MDA2 = Molecular Staging REPLI-g™ 2500S, FLP = Rubicon OmniPlex™.
[b]gDNA tissue source: A total of 20 gDNA samples were extracted from each tissue type and used as template for each WGA protocol.
[c]No template controls (no gDNA input into the WGA) from the MDA1 (N = 3), MDA2 (N = 3), and FLP (N = 1) protocols produced 3.38 (mean), 21.12 (mean), and 0.41 μg of total wgaDNA (55%, 103%, and 5% of median total wgaDNA produced with gDNA samples), with 56%, 15%, and 80% of the wgaDNA produced as ssDNA, respectively.
[d]wgaDNA yield consists of the total sum of estimated double-stranded and single-stranded DNA wgaDNA as described in Materials and Methods.
[e]ssDNA % is the proportion of wgaDNA estimated to be present as single-stranded DNA as described in Materials and Methods.
[f]$P < 0.001$, differences between each WGA protocol with respect to the fraction of single-stranded wgaDNA, Wilcoxon rank sum test.
[g]$P = 0.0068, 0.0284$, and 0.0026, overall differences in FLP %ssDNA, mouthwash vs. lymphoblast %ssDNA, buffy coat vs. lymphoblast %ssDNA, respectively, Wilcoxon rank sum test.
[h]$P = 0.02$, buffy coat vs. lymphoblast FLP wgaDNA yield, Wilcoxon rank sum test.

TABLE 2. AmpF*l*STR® Identifiler® Assay, No Amplification, and Discordant Genotypes by Relative Fluorescence Units (RFU)

				Height of allele >50 RFU		Height of allele >100 RFU		Height of allele >150 RFU		Height of allele >200 RFU		Height of allele >250 RFU	
DNA type	gDNA sample source	N (sample)	N (attempted genotype)	N (no amp.)	N (discordant)	N (no amp.)	N (discordant)	N (no amp.)	N (discordant)	N (no amp.)	N (discordant)	N (no amp.)	N (discordant)
Genomic DNA[a]	Mouthwash	20	640	0	0	0	0	0	0	2	0	7	0
	Buffy coat	20	640	0	0	3	0	5	0	11	0	17	0
	Lymphocyte	34	1088	0	0	2	0	4	0	15	0	38	0
	Total	74	2368	0	0	5	0	9	0	28	0	62	0.00
MDA1[b]	Mouthwash	20	640	19	45	50	36	89	32	137	30	173	27
	Buffy coat	20	640[c]	3	50	31	39	57	36	97	35	152	30
	Lymphocyte	20	640	29	38	51	32	73	26	129	18	172	17
	Total	60	1920	51	133	132	107	219	94	363	83	497	74
MDA2[b]	Mouthwash	20	640	19	62	54	53	93	45	123	43	170	42
	Buffy coat	20	640[c]	5	50	35	44	68	41	105	39	143	36
	Lymphocyte	20	640	10	38	40	31	71	27	105	21[e]	154	20
	Total	60	1920	34	150	129	128	232	113	333	103	467	98
FLP	Mouthwash	20	640	95	134	205	67	265	41	296	33	328	27
	Buffy coat	19[d]	608[c]	63	148	155	85	206	61	247	41	317	33
	Lymphocyte	20	640	84	148	196	75	237	57	286	38	275	31
	Total	59	1888	242	430	556	227	708	159	829	112	920	91

[a]gDNA exhibited significantly greater genotype completion and concordance rates compared to all WGA protocols and wgaDNA tissue sources at all RFU thresholds except for MDA1 buffy coat wgaDNA (P values < 0.001), Chi-square test.
[b]The MDA protocols resulted in significantly higher completion ($P < 0.001$) and concordance ($P < 0.05$) rates compared to the FLP protocol at all RFU thresholds, Chi-square test.
[c]$P < 0.001$ over all WGA methods, $P < 0.001$ for MDA1, $P = 0.01$ for MDA2, $P = 0.06$ for FLP, buffy coat vs. mouthwash or lymphoblast wgaDNA, STR genotype completion rate, 50 RFUs, Chi-square test.
[d]One buffy coat gDNA yielded substantially reduced wgaDNA after FLP, therefore only 19 FLP wgaDNAs were genotyped.
[e]$P = 0.001$, lymphoblast vs. mouthwash MDA2 wgaDNA STR discordance rate, Chi-square test.

TABLE 3. Summary of gDNA and wgaDNA SNP Genotyping Data (N = 49 TaqMan® SNP Assays)

DNA type	gDNA sample source	N (sample)	N (attempted genotype)	N (completed genotype)	% completion	N (undeter.)	% undeter.	N (no amp.)	% no amp.	N (discordant)	% discordance
Genomic DNA	Mouthwash	20	1960	1926	98.27	19	0.97	15	0.77	0	0.00
	Buffy coat	20	1960	1917	97.81	36	1.84	7	0.36	0	0.00
	Lymphocyte	34	3332	3293	98.83	35	1.05	4	0.12	4	0.12
	Total	74	7252	7136	98.40[a]	90	1.24[a]	26	0.36	4	0.06
MDA1	Mouthwash	20	1960	1889	96.38	57	2.91	14	0.71	13[b]	0.69
	Buffy coat	20	1960	1884	96.12	71	3.62	5	0.26	1	0.05
	Lymphocyte	20	1960	1879	95.87	70	3.57	11	0.56	4	0.21
	Total	60	5880	5652	96.12[a]	198	3.37[a]	30	0.51	18	0.32
MDA2	Mouthwash	20	1960	1854	94.59	103	5.26	3	0.15	24	1.29
	Buffy coat	20	1960	1862	95.00	85	4.34	13	0.66	12	0.64
	Lymphocyte	20	1960	1908	97.35	49[c]	2.50	3	0.15	6[c]	0.32
	Total	60	5880	5624	95.65[a]	237	4.03[a]	19	0.32	42	0.75
FLP	Mouthwash	20	1960	1811	92.40	140	7.14	9	0.46	3	0.17
	Buffy coat	19	1862	1742	93.56	114	6.12	6	0.32	0	0.00
	Lymphocyte	20	1960	1807	92.19	145	7.40	8	0.41	5	0.28
	Total	59	5782	5360	92.70[a]	399	6.90[a]	23	0.40	8[d]	0.15[e]

[a]$P < 0.001$, completion and undetermined rates, gDNA vs. MDA and FLP, and MDA vs. FLP.
[b]$P = 0.002$, discordant SNP genotyping rate, MDA1 mouthwash wgaDNA vs. MDA1 lymphoblast and buffy coat wgaDNA, Chi-square test.
[c]$P < 0.001$ and $P = 0.002$, undetermined and discordant SNP genotyping rate, respectively, MDA2 lymphoblast wgaDNA vs. MDA2 mouthwash and buffy coat wgaDNA, Chi-square test.
[d]$P < 0.001$, discordant SNP genotyping rate FLP vs. MDA2, Chi-square test.
[e]$P < 0.001$, discordance rate, FLP versus MDA2.

genotypes, when manually scored, were <20% for the MDA and the FLP protocols (data not shown).

SNP Genotyping With the TaqMan Assay (N = 49 SNPs)

TaqMan SNP genotypes from gDNA and wgaDNA were compared to evaluate TaqMan SNP assay performance (Table 3). gDNA exhibited completion, undetermined, and discordant genotype rates within the range of published error for TaqMan assays [Ranade et al., 2001]. gDNA exhibited significantly higher completion and significantly lower undetermined SNP genotype rates compared to wgaDNA from all three WGA protocols. Rates for no amplification for all three WGA protocols were similar to those observed in the gDNA samples. FLP wgaDNA exhibited the lowest discordance rate of the WGA protocols and significantly lower than the MDA2 protocol. The three WGA protocols had overall SNP genotype concordance rates of >99.1%, compared to nearly complete concordance between gDNA replicates (99.94%). For all three protocols, there was no instance in which a gDNA homozygote for one allele was scored as a homozygote for the other allele when genotyping wgaDNA, or vice versa (Supplementary Table S4). Most wgaDNA genotype discordances involved the heterozygote genotype cluster and one of the homozygote clusters, and the majority of the wgaDNA discordant and undetermined genotype failures occurred when a gDNA heterozygote genotype was scored as a wgaDNA homozygote genotype. In addition, no SNP assay completely failed for the MDA protocols, suggesting that the appropriate sequence surrounding the 49 SNP loci was amplified. However, two TaqMan SNP assays (*CDKN1A* and *CDKN2A*) had very high rates of undetermined SNP genotypes with FLP (Supplementary Table S5), suggesting, similar to the results observed with STR genotyping, that the FLP protocol exhibits modest locus bias.

%GC and SNP Genotyping Performance

We investigated whether local sequence context, specifically %GC sequence content, might affect the performance of the WGA reaction and/or the TaqMan genotyping assay (Supplementary Table S5). We used both regional (5 kb 5′ and 3′ of the SNP) and TaqMan genotyping assay amplicon %GC measures in correlation analysis of SNP genotype failures, in order to test for independence of each amplification reaction (WGA and TaqMan, respectively) and %GC. However, note that regional %GC (mean: 44.9%; median: 43.3%; range: 33.8–65.8%) is significantly correlated (Spearman $R = 0.654$; $P < 0.001$) with amplicon %GC. gDNA samples exhibited a modest statistically significant positive correlation between regional %GC and the small number of discordant genotypes ($P = 0.028$ and a trend with amplicon %GC), but no significant correlation was observed with the much larger number of undetermined genotypes. MDA wgaDNA undetermined and discordant TaqMan genotypes exhibited modest statistically significant correlations with %GC, but the pattern of correlation differed between the MDA protocols. MDA1 wgaDNA exhibited a significant positive correlation ($P = 0.027$) between undetermined genotypes and TaqMan amplicon %GC, while MDA2 wgaDNA exhibited a significant positive correlation ($P = 0.040$) between discordant genotypes and regional %GC. The FLP protocol exhibited both a modest significant positive correlation ($P = 0.04$) between discordant genotypes and regional %GC, and stronger significant positive correlations ($P < 0.001$) between undetermined genotypes and both regional and TaqMan amplicon %GC. Approximately 7% and

19% of the TaqMan genotyping performance bias exhibited by MDA and FLP wgaDNA, respectively, can be attributed to %GC.

Effects of gDNA Source on wgaDNA

We investigated whether the yield, composition, or genotyping performance of wgaDNA differed by the three tissue sources within the WGA protocols used in this study. First, FLP wgaDNA exhibited increased yield from buffy coat gDNA samples and the composition of the FLP wgaDNA varied significantly overall between the three tissue sources, between mouthwash and lymphoblast, and between buffy coat and lymphoblast (Table 1). Second, buffy coat wgaDNA exhibited significantly higher STR genotype completion rates than mouthwash or lymphoblast wgaDNA for all WGA methods, and lymphoblast MDA2 wgaDNA had a significantly lower discordance rate than mouthwash MDA2 wgaDNA (Table 2). However, there were no significant differences by individual STR locus between gDNA tissue sources with respect to discordant or no amplification genotype failures. Finally, there were three significant pairwise differences in SNP genotyping performance by wgaDNA gDNA source (Table 3): MDA2 lymphoblast wgaDNA exhibited significantly lower undetermined and discordant SNP genotyping rates compared to MDA2 mouthwash and buffy coat wgaDNA, and MDA1 mouthwash wgaDNA exhibited a significantly higher discordant SNP genotyping rate compared to lymphoblast and buffy coat MDA1 wgaDNA.

Results of Analysis of wgaDNA Derived from NTC Samples

We observed that NTCs included in the WGA produced substantial amounts of both ssDNA and dsDNA in all three protocols (Table 1). The proportion of ssDNA in the wgaDNA derived from NTCs was much greater than the proportion of ssDNA in the wgaDNA derived from gDNA for the MDA1 protocol, but similar for the other WGA protocols (Table 1). The AmpF*l*STR Identifiler assay, which requires PCR amplification and fragment analysis of specific PCR products, exhibited nine false-positive STR genotypes at seven STR loci in the seven NTCs analyzed, though most (92.6%) NTC STR genotype attempts yielded no genotypes (data not shown). These NTC genotypes were mostly homozygote genotypes (78%), and were of low signal strength (mean = 138 RFUs). We observed that fluorescent signals from MDA NTCs were included in a TaqMan SNP genotype cluster more often ($P = 0.01$ for MDA1; $P = 0.07$ for MDA2) than observed with NTCs introduced into the TaqMan SNP assay as negative controls (10.9% and 3.4% vs. 1.7%, respectively). As with the NTC STR genotypes, the NTC SNP genotypes were not common (92.9% of NTC SNP genotype attempts yield no genotypes), and were nearly all homozygote genotypes (97.7%). The FLP protocol did not exhibit increased numbers of false-positive genotypes with the NTC included in this WGA protocol. No significant differences were observed with respect to which labeled TaqMan allelic probe exhibited the increased fluorescence, suggesting nonspecific probe degradation (data not shown).

Statistical Power

The statistical power of this study is sufficiently robust, given the number of attempted AmpF*l*STR Identifiler STR and TaqMan SNP genotypes, to detect significant differences in completion, undetermined, and concordance rates by WGA protocol, tissue type, and genotyping application. For example, the power to detect a 2% difference in AmpF*l*STR Identifiler STR genotype completion rates, e.g., between 96% and 98%, with the number of attempted AmpF*l*STR Identifiler STR genotypes (N = 1,920) for the two MDA protocols, is 95%. The power to detect a 2% difference in AmpF*l*STR Identifiler STR genotype discordance rates, e.g., between 6% and 8%, with the number of attempted AmpF*l*STR Identifiler genotypes (N = 1,920) for the two MDA protocols, is 68%. The power to detect a 1.5% difference in TaqMan SNP assay undetermined genotype rates by tissue type within WGA protocol, e.g., between 2.5% and 4%, given the number of TaqMan SNP attempted genotypes by tissue type within WGA protocol (N = 1,960), is 76%. The power to detect similar differences in genotyping performance between tissue types within WGA protocols was lower, yet the power to detect substantial differences in such strata, e.g., a difference in genotyping rates between 75% and 95%, was robust at 74%.

DISCUSSION

WGA promises to reduce the concern that available quantities of gDNA may become limiting for molecular epidemiologic, clinical diagnostic, and forensic studies. We have evaluated two methods of WGA (MDA and FLP), with respect to utility for two widely used genotyping applications (STR profiling and SNP genotyping), using gDNA derived from three tissue sources. In this comparative evaluation, MDA wgaDNA exhibited a significantly greater proportion of double-stranded wgaDNA, significantly greater STR and SNP genotype completion rates, and significantly reduced STR genotype discordance rates and sensitivity to %GC content, than did FLP wgaDNA.

Significant reductions in STR completion and concordance rates for FLP vs. MDA2 wgaDNA derived from lymphoblast gDNA have been observed in a comparative evaluation of both methods [Bark et al., 2004]. The same study also observed a significantly increased rate of unbalanced heterozygotes in FLP wgaDNA, similar to the significant differences exhibited between FLP and MDA PHRs and GQ scores in this study. Increased STR allelic dropout and greater asymmetry in peak area ratios has been observed in low copy number templates, i.e., 12 or 25 pg of gDNA, PCR-amplified with the AmpF*l*STR Second Generation Multiplex (SGM) Plus® STR panel (10 of the 15 STR loci used in the AmpF*l*STR Identifiler panel) [Whitaker et al., 2001]. This suggests that the reduced STR genotyping performance with wgaDNA observed in this study, and in Bark et al. [2004], is due to stochastic WGA and subsequent PCR amplification at repetitive loci.

TaqMan SNP genotyping assay completion rates have previously been reported to be reduced by ~2% for MDA1 wgaDNA derived from whole blood and somewhat greater when derived from buccal cells [Tranah et al., 2003]. We observed similar reductions in SNP genotyping completion rates with both MDA protocols with wgaDNA derived from gDNA extracted from three tissue sources. We observed that TaqMan assay fluorescence scatter associated with each genotype cluster is often visibly greater with wgaDNA samples than from gDNA samples, leading to data points falling outside operator-defined genotype clusters, i.e., resulting in more undetermined genotypes, especially with the FLP protocol. We also observed that the majority of discordant TaqMan SNP genotypes from both MDA and FLP wgaDNA are due to allelic dropout from individuals genotyped as gDNA heterozygotes, and that wgaDNA SNP genotyping performance is significantly correlated with %GC content. In order to reduce the stochastic amplification observed at both repetitive and nonrepetitive loci upon STR and SNP genotyping of wgaDNA, it may be necessary to use increased amounts of gDNA as template for WGA, to use WGA methods

that result in reduced amounts of ssDNA, to consistently denature gDNA samples prior to WGA, and to avoid genotyping repetitive loci and/or loci with increased %GC.

For example, in an evaluation of MDA1 WGA using laser capture microdissection to isolate 100, 300, 750, 1,500, and 3,000 cells (~0.7–21 ng of genomic DNA), and the same genotyping methods as in this study, genotyping completion and concordance rates increased when using DNA extracted from more cells, or when using wgaDNA pooled from multiple WGA reactions [Rook et al., 2004]. Similarly, in an analysis of a range of gDNA inputs into MDA1 WGA (0.003, 0.03, 0.3, and 3 ng) and using a four-color minisequencing microarray-based SNP genotyping method, MDA1 wgaDNA exhibited acceptable genotyping success rates (completion rate and genotype signal ratios) only when 3 ng of gDNA was used as template [Lovmar et al., 2003]. The quantity of input gDNA used in the WGA protocols in this study (4 ng for MDA and 5 ng for FLP) had been thought to be sufficient to generate unbiased WGA of the genome [Dean et al., 2002; Hosono et al., 2003; Lovmar et al., 2003], however, Lasken and Egholm [2003] have recently recommended that a minimum of 10 ng of undegraded gDNA should be used to avoid stochastic amplification of alleles during the WGA reaction. Such recommendations suggest that the routine use of WGA in studies with highly limiting quantities of gDNA input, such as studies of archeological remains [Handt et al., 1996], prenatal diagnostic procedures [Paunio et al., 1996], analysis of pathological specimens [Rook et al., 2004], and forensic studies [Sorensen et al., 2004], may risk stochastic amplification of loci or alleles. However, it should be possible to distinguish false positive wgaDNA genotypes derived from NTCs by their low completion and heterozygosity rate from true positive wgaDNA genotype, especially when working with loci for which heterozygosity is characterized, based on the characteristics of the NTC false-positive STR and SNP genotypes observed in this study.

With respect to the related issues of amounts of template gDNA, denaturation of template gDNA prior to WGA, and %GC content, a significantly reduced SNP genotyping completion rate has been reported [Paez et al., 2004] with MDA2 wgaDNA when gDNA was not denatured (88.92% for MDA2 wgaDNA derived from undenatured gDNA, 92.06% from MDA2 wgaDNA derived from denatured gDNA, and 92.93% from gDNA, all using 10 ng of DNA), using the Affymetrix (Santa Clara, CA; www.affymetrix.com) two-color variant detection array 10 K SNP genotyping panel [Matsuzaki et al., 2004]. Similarly, a significantly reduced wgaDNA SNP genotyping concordance rate (to gDNA SNP genotypes) of 99.59% was observed when gDNA samples were not denatured prior to MDA2 WGA, compared to a wgaDNA SNP genotyping concordance rate of 99.80% observed when the gDNA template was denatured prior to MDA2 [Paez et al., 2004]. Interestingly, we observed no significant differences in MDA1 and MDA2 wgaDNA STR or SNP genotype completion or discordance rates, even though the two MDA protocols differed with respect to denaturation of gDNA prior to WGA in this evaluation (MDA1 was performed with no denaturation, MDA2 had chemical denaturation). However, our sample of attempted genotypes was not large enough to detect a statistically significant difference of 0.2% in genotyping completion or concordance rate. In a study using the Sentrix™ Array SNP genotyping platform (Illumina, San Diego, CA; www.illumina.com), MDA2 and FLP wgaDNA genotypes have been reported to be 99.86% and 99.89% concordant with gDNA genotypes, respectively [Barker et al., 2004]. While the amount of gDNA used as template for WGA in Barker et al. [2004] was greater than the amount of gDNA used as template for WGA in this study (2 × for MDA2 and 25 × for FLP), the high concordance rate observed between wgaDNA and gDNA SNP genotyping using the Illumina linkage panel suggests that SNP genotyping assays that use a restricted range of %GC sequence content coupled with multiple allelic discrimination steps, such as the Golden Gate™ Assay [Fan et al., 2003], may be less susceptible to the effects of %GC content when genotyping wgaDNA than genotyping methods with fewer allelic discrimination steps.

We conclude that further work is required to define the necessary range of input gDNA for WGA protocols for subsequent robust wgaDNA genotyping performance. It may be necessary to ensure that sufficient gDNA remains available for validation of findings if the source of gDNA for WGA is limiting. If such gDNA is not available, then qualification of wgaDNA analysis results based on an understanding of the error associated with various molecular analyses commonly performed in molecular epidemiology, clinical diagnostics, and forensic DNA analysis, may be necessary.

ACKNOWLEDGMENTS

We thank Meredith Yeager, Ph.D. for SNP annotation and validation and Vinita Puri for assistance with SNP genotyping.

REFERENCES

Bark C, Pettengill J, Tsai YBP, Golembieski J, Gearhart J, Stewart L, Zilka M, Doheny K. 2004. Performance of whole genome amplified samples for microsatellite genotyping. 54th Annual Meeting of the American Society of Human Genetics, October 26–29, 2004, Toronto, Ontario, Canada. Abstract 1957.

Barker DL, Hansen MS, Faruqi AF, Giannola D, Irsula OR, Lasken RS, Latterich M, Makarov V, Oliphant A, Pinter JH, Shen R, Sleptsova I, Ziehler W, Lai E. 2004. Two methods of whole-genome amplification enable accurate genotyping across a 2320-SNP linkage panel. Genome Res 14:901–907.

Blanco L, Bernad A, Lazaro JM, Martin G, Garmendia C, Salas M. 1989. Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. J Biol Chem 264:8935–8940.

Cheung VG, Nelson SF. 1996. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. Proc Natl Acad Sci USA 93:14676–14679.

Dean FB, Nelson JR, Giesler TL, Lasken RS. 2001. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. Genome Res 11:1095–1099.

Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore SF, Egholm M, Lasken RS. 2002. Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci USA 99:5261–5266.

Dietmaier W, Hartmann A, Wallinger S, Heinmoller E, Kerner T, Endl E, Jauch KW, Hofstadter F, Ruschoff J. 1999. Multiple mutation analyses in single tumor cells with improved whole genome amplification. Am J Pathol 154:83–95.

Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F, Gunderson KL, Hanse M, Steemers F, Butler SL, Deloukas P, Galver L, Hunt S, Mcbride C, Bibikova M, Rubano T, Chan J, Wickham E, Doucet D, Chang W, Campbell D, Zhang B, Kruglyak S, Bentley D, Haas J, Rigault P, Zhou L, Stuelpnagel J, Chee MS. 2003.

Highly parallel SNP genotyping. Cold Spring Harb Symp Quant Biol 68:69–78.

Garcia-Closas M, Egan KM, Abruzzo J, Newcomb PA, Titus-Ernstoff L, Franklin T, Bender PK, Beck JC, Le Marchand L, Lum A, Alavanja M, Hayes RB, Rutter J, Buetow K, Brinton LA, Rothman N. 2001. Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. Cancer Epidemiol Biomarkers Prev 10:687–696.

Handt O, Krings M, Ward RH, Paabo S. 1996. The retrieval of ancient human DNA sequences. Am J Hum Genet 59: 368–376.

Haque KA, Pfeiffer RM, Beerman MB, Struewing JP, Chanock SJ, Bergen AW. 2003. Performance of high-throughput DNA quantification methods. BMC Biotechnol 3:20.

Hosono S, Faruqi AF, Dean FB, Du Y, Sun Z, Wu X, Du J, Kingsmore SF, Egholm M, Lasken RS. 2003. Unbiased whole-genome amplification directly from clinical samples. Genome Res 13:954–964.

Klein CA, Schmidt-Kittler O, Schardt JA, Pantel K, Speicher MR, Riethmuller G. 1999. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. Proc Natl Acad Sci USA 96:4494–4499.

Langmore JP. 2002. Rubicon Genomics, Inc. Pharmacogenomics 3:557–560.

Lasken RS, Egholm M. 2003. Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. Trends Biotechnol 21:531–535.

Lizardi PM. 2000. Multiple displacement amplification. U.S. Patent 6,124,120.

Lovmar L, Fredriksson M, Liljedahl U, Sigurdsson S, Syvanen AC. 2003. Quantitative evaluation by minisequencing and micro-arrays reveals accurate multiplexed SNP genotyping of whole genome amplified DNA. Nucleic Acids Res 31:e129.

Lucito R, Nakimura M, West JA, Han Y, Chin K, Jensen K, Mccombie R, Gray JW, Wigler M. 1998. Genetic analysis using genomic representations. Proc Natl Acad Sci USA 95: 4487–4492.

Matsuzaki H, Loi H, Dong S, Tsai YY, Fang J, Law J, Di X, Liu WM, Yang G, Liu G, Huang J, Kennedy GC, Ryder TB, Marcus GA, Walsh PS, Shriver MD, Puck JM, Jones KW, Mei R. 2004. Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. Genome Res 14: 414–425.

Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Brown AL, Budowle B. 2001. Validation of STR typing by capillary electrophoresis. J Forensic Sci 46:661–676.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb Symp Quant Biol 51:263–273.

Packer BR, Yeager M, Staats B, Welch R, Crenshaw A, Kiley M, Eckert A, Beerman M, Miller E, Bergen A, Rothman N, Strausberg R, Chanock SJ. 2004. SNP500Cancer: a public resource for sequence validation and assay development for genetic variation in candidate genes. Nucleic Acids Res 32: D528–D532.

Paez JG, Lin M, Beroukhim R, Lee JC, Zhao X, Richter DJ, Gabriel S, Herman P, Sasaki H, Altshuler D, Li C, Meyerson M, Sellers WR. 2004. Genome coverage and sequence fidelity of phi29 polymerase-based multiple strand displacement whole genome amplification. Nucleic Acids Res 32:e71.

Paunio T, Reima I, Syvanen AC. 1996. Preimplantation diagnosis by whole-genome amplification, PCR amplification, and solid-phase minisequencing of blastomere DNA. Clin Chem 42: 1382–1390.

Ranade K, Chang MS, Ting CT, Pei D, Hsiao CF, Olivier M, Pesich R, Hebert J, Chen YD, Dzau VJ, Curb D, Olshen R, Risch N, Cox DR, Botstein D. 2001. High-throughput genotyping with single nucleotide polymorphisms. Genome Res 11:1262–1268.

Rook MS, Delach SM, Deyneko G, Worlock A, Wolfe JL. 2004. Whole genome amplification of DNA from laser capture-microdissected tissue for high-throughput single nucleotide polymorphism and short tandem repeat genotyping. Am J Pathol 164:23–33.

Sorensen KJ, Turteltaub K, Vrankovich G, Williams J, Christian AT. 2004. Whole-genome amplification of DNA from residual cells left by incidental contact. Anal Biochem 324:312–314.

Tanabe C, Aoyagi K, Sakiyama T, Kohno T, Yanagitani N, Akimoto S, Sakamoto M, Sakamoto H, Yokota J, Ohki M, Terada M, Yoshida T, Sasaki H. 2003. Evaluation of a whole-genome amplification method based on adaptor-ligation PCR of randomly sheared genomic DNA. Genes Chromosomes Cancer 38:168–176.

Telenius H, Carter N P, Bebb CE, Nordenskjold M, Ponder BA, Tunnacliffe A. 1992. Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. Genomics 13:718–725.

Tranah GJ, Lescault PJ, Hunter DJ, De Vivo I. 2003. Multiple displacement amplification prior to single nucleotide polymorphism genotyping in epidemiologic studies. Biotechnol Lett 25: 1031–1036.

Wells D, Sherlock JK, Handyside AH, Delhanty JD. 1999. Detailed chromosomal and molecular genetic analysis of single cells by whole genome amplification and comparative genomic hybridisation. Nucleic Acids Res 27:1214–1218.

Whitaker JP, Cotton EA, Gill P. 2001. A comparison of the characteristics of profiles produced with the AMPFlSTR SGM Plus multiplex system for both standard and low copy number (LCN) STR DNA analysis. Forensic Sci Int 123: 215–223.

Zhang L, Cui X, Schmitt K, Hubert R, Navidi W, Arnheim N. 1992. Whole genome amplification from a single cell: implications for genetic analysis. Proc Natl Acad Sci USA 89:5847–5851.